JBC Papers in Press. Published on June 20, 2012 as Manuscript M112.382960
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M112.382960

Impact of Protein-Protein Interactions on Global Intermolecular Translocation Rates of the Transcription Factors Sox2 and Oct1 Between DNA Cognate Sites Analyzed by *z*-Exchange NMR Spectroscopy

Yuki Takayama and G. Marius Clore[1]

From the Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520

Running title: Impact of Sox2/Oct1 interaction on intermolecular translocation

[1]Address correspondence to: G. Marius Clore, Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520. Tel.: 301-496-0788; Fax: 301-496-0825. E-mail: mariusc@intra.niddk.nih.gov.

Key words: Protein-DNA interactions / translocation / jumping and intersegment transfer / *z*-exchange spectroscopy

Background: Sox2 and Oct1 interact on a variety of promoters to regulate transcription.
Results: Global intermolecular translocation rates in a ternary Sox2·Oct1·*FGF4*-DNA complex have been analyzed by *z*-exchange spectroscopy.
Conclusion: Translocation is modulated by protein-protein interactions on the DNA.
Significance: The data suggest a model for the sequence of binding events involved in combinatorial control of gene regulation by Sox2 and Oct1.

Oct1 and Sox2 synergistically regulate developmental genes by binding to adjacent sites within promoters. We have investigated the kinetics of global intermolecular translocation of Sox2 and Oct1 between cognate sites located on different DNA molecules by *z*-exchange NMR spectroscopy. In the *Hoxb1* promoter the Sox2 and Oct1 sites are immediately adjacent to one another, and the intermolecular translocation rates are too slow to be measured by *z*-exchange spectroscopy. By introducing a three base pair insertion between the Sox2 and Oct1 sites to mimic the spacing in the *FGF4* enhancer, the interprotein contact surface is reduced and the translocation rates are increased. Interaction between Sox2 and the POU_S domain of Oct1 does not affect the translocation mechanism but modulates the rates. Translocation involves only jumping (dissociation and reassociation) for Sox2, but both jumping and direct intersegment transfer (no dissociation into free solution) for Oct1. The dissociation (k_{off} ~ 1.5 s^{-1}) and association (k_{on} ~ 5.1x10^9 $M^{-1}s^{-1}$) rate constants for Sox2 are reduced four-fold and increased five-fold, respectively, in the presence of Oct1. k_{off} (~3.5 s^{-1}) for Oct1 is unaffected by Sox2, while k_{on} (~1.3x10^9 $M^{-1}s^{-1}$) is increased ~13-fold. The direct intermolecular translocation rate (k_{inter} ~1.8x10^4 $M^{-1}s^{-1}$) for the POU_S domain of Oct1 is reduced two-fold by Sox2, while that for the POU_{HD} domain of Oct1 (k_{inter} ~ 1.7x10^4 $M^{-1}s^{-1}$) remains unaltered, consistent with the absence of contacts between Sox2 and POU_{HD}. The data suggest a model for the sequence of binding events involved in synergistic gene regulation by Sox2 and Oct1.

In eukaryotes, combinatorial control of gene expression involves the formation of multi-transcription factor complexes that effectively integrate a wide range of signaling pathways to provide temporal and cell specific transcription regulation (1). An example of this phenomenon is provided by members of the Sox and Oct transcription factor families that interact with a variety of DNA promoter/enhancer elements to regulate transcription during embryogenesis and neural development (2,3). Sox2 is a member of the HMG-box family of architectural factors that bind to the minor groove of DNA and bend the DNA by 50-90° (4). Oct1 comprises two major-groove DNA binding domains, a POU specific domain (POU_S) and a homeodomain (POU_{HD}), connected by a flexible linker (5,6). Structures of

ternary complexes of Sox2 and Oct1 bound to regulatory elements within the *Hoxb1* promoter (7) and *fibroblast growth factor-4* (*FGF4*) enhancer (8), differing in the spacing between the Sox2 and Oct1 binding sites, have also been solved by NMR and crystallography, respectively. While three-dimensional structures of these binary and ternary protein-DNA complexes have yielded a wealth of static information regarding the structural basis of protein-DNA recognition by Sox2 and Oct1, less is known of the mechanisms whereby these transcription factors locate their specific target sites within an overwhelming sea of non-specific DNA (9-11), especially within the context of multi-transcription factor complexes.

Recently we have made use of NMR paramagnetic relaxation enhancement measurements (12,13) to detect and characterize transient sparsely-populated, spectroscopically invisible states of protein-DNA complexes that are critical to the target search process (14-19). In the context of a specific complex, these intermediate states, which occupy non-specific DNA sites and have lifetimes of less than 250-500 μs, are populated at less than 0.5% and are involved in both one-dimensional rotation-coupled sliding along the DNA and direct intersegment transfer from one DNA molecule to another. This methodology has been used to study target searching by the homeodomain transcription factor HoxD9 (15,16), the bi-domain transcription factor Oct1 (17), the minor groove binding architectural factor Sox2 (18), and a ternary complex of Oct1 and Sox2 bound to the regulatory element within the *Hoxb1* promoter (18).

In addition to rapid translocation events involving sparsely-populated states, global intermolecular translocation of the major spectroscopically visible species between specific sites on different DNA molecules occurs on a much slower time scale (10 ms to 1 s) and can be directly observed and kinetically analyzed using two-dimensional *z*-exchange NMR spectroscopy (18-21). In the case of HoxD9, global intermolecular translocation occurs exclusively by direct intersegment transfer without necessitating dissociation of the protein into free solution (20). For Sox2, on the other hand, global intermolecular translocation between specific DNA sites proceeds entirely by jumping, a process that entails complete dissociation of Sox2 from the DNA into free solution, followed by reassociation (18). Direct intersegment transfer and, to a lesser extent, jumping occur with Oct1 (21). When Oct1 and Sox2 form a ternary complex on the *Hoxb1* promoter, the translocation rate between specific DNA sites is reduced by over an order of magnitude and can no longer be studied by *z*-exchange spectroscopy (18).

To study the mechanism and kinetics of global intermolecular translocation of the protein components within an Oct1·Sox2·DNA ternary complex, we therefore chose to focus on the *FGF4* enhancer, where the spacing between the Oct1 and Sox2 recognition sites is increased by 3 base pairs (22) relative to that within the *Hoxb1* regulatory element (23) (Fig. 1A). The interaction surface between Oct1 and Sox2 in the ternary complex on the *FGF4* enhancer (8) is altered and reduced relative to that on *Hoxb1* (7), and as a result the strength of the interaction between the two proteins is weakened and the rate of translocation is increased sufficiently to permit the application of *z*-exchange spectroscopy. Here we show how protein-protein interactions between Oct1 and Sox2 on the *FGF4* promoter modulate the kinetics of global intermolecular translocation.

Experimental Procedures

Sample preparation — The POU region (POU_S + POU_{HD}) of human Oct1 (residues 280-442) and the HMG-box domain of Sox2 (residues 38-121) were expressed and purified as described previously (7,18). Uniform $^{2}H/^{15}N$ isotopic labeling was achieved by growing *Escherichia coli* BL21-CodonPlus(DE3)-RIPL cells in minimal medium with 99.9% D_2O, d_7-glucose and $^{15}NH_4Cl$. Single-stranded unmodified and rhodamine-conjugated DNA oligonucleotides were purchased from Invitrogen Inc. and Midland Certified Reagents, respectively, and purified by an anion-exchange chromatography on a Mono-Q (GE Healthcare Bioscience) column with a NaCl gradient in a buffer of 50 mM Tris-HCl, pH 7.5 and 1 mM EDTA. After annealing, DNA duplexes were further purified by anion-exchange chromatography to remove any residual single-stranded DNA (24). Fluorescence anisotropy and NMR samples were prepared in 10 mM PIPES, 150 mM NaCl, 94% H_2O/6% D_2O, pH 6.5.

Fluorescence anisotropy — The K_D for the binding of Sox2 and Oct1 to *FGF4*-DNA and of Oct1 to the Sox2·*FGF4*-DNA complex at 30°C

was determined by fluorescence anisotropy using a Jobin-Yvon FluoroMax-3 spectrometer as described previously (16). The wavelengths for excitation and emission were 550 and 580 nm, respectively. Sox2 (0 to 154 nM) and Oct1 (0 to 351 nM) were added to 1.5 and 10 nM rhodamine-conjugated 32-bp *FGF4*-DNA duplex, respectively. The K_D was calculated from the titration data as described previously (14). Since the difference in fluorescence anisotropy for the binding of Sox2 to the Oct1·*FGF4*-DNA complex is too small to permit an accurate K_D determination, the K_D for Oct1 binding to *FGF4*-DNA (10 nM) in the presence of Sox2 (100 nM) was measured. Under these conditions, ~95% of the DNA is present as a specific Oct1-complex and <3% is present as a non-specific complex. The effect of Oct1 on the K_D for Sox2 binding was then determined from a thermodynamic cycle.

NMR spectroscopy — All NMR experiments were carried out at 303 K on Bruker 600 MHz spectrometers equipped with *z*-gradient triple resonance cryoprobes. Spectra were processed using NMRPipe (25) and analyzed using the program NMRView (26).

Exchange rates were measured using transverse relaxation optimized (TROSY)-based *z*-exchange spectroscopy (27) with at least eight different mixing times between 20 and 600 ms. Fitting the time dependence of the exchange and auto peaks to derive kinetic rate constants was as described in (20,21).

Results and Discussion

Interaction of Sox2 and Oct1 on the FGF4-promoter — In *Hoxb1*-DNA promoter, the Sox2 and Oct1 binding sites are immediately adjacent to one another (7,23), while there is 3-bp insertion between the Sox2 and Oct1 cognate sites in the *FGF4*-enhancer (8,22). The sequence of the 32 bp *FGF4*-like DNA duplex (hereafter referred to as "*FGF4*-DNA") containing the specific binding sites for Sox2 and Oct1 is shown in Fig. 1A (left). This sequence does not represent the actual sequence within the *FGF4* enhancer, but rather simply adds the three base pair insertion (TGG) between the Sox2 and Oct1 binding sites found in the *FGF4* enhancer to the *Hoxb1* promoter sequence (Fig. 1A, right). The sequences of the Sox2 and Oct1 specific sites as well as the sequences on the 5' end of the Sox2 and 3' end of the Oct1 binding sites are thus identical to the *Hoxb1* promoter sequence. This ensures that differences in equilibrium dissociation constants, and rates of intermolecular translocation between the "*FGF4* enhancer" and *Hoxb1* promoter DNA duplexes reflect only the impact of the three base pair insertion between the Sox2 and Oct1 binding sites.

The different spacing of the Sox2 and Oct1 specific sites on the *FGF4*-DNA and *Hoxb1*-DNA duplexes alters the relative orientations of the two proteins and the protein-protein interface in the two ternary complexes (7,8). In the *Hoxb1* ternary complex, the protein-protein interface is formed between residues Lys59 to Lys73 (helix 3) of Sox2 and residues Lys14 to Thr26 (helix 1) of the POU_S domain of Oct1 (7). The protein-protein interface on the *FGF4* enhancer, on the other hand, involves only two residues from Sox2 (Arg81 and Arg82) and four residues from the POU_S domain (Ile25, Gly28, Thr30, and Asp33) (8). This is in complete agreement with the location and breadth of the corresponding $^1H_N/^{15}N$ chemical shift perturbation profiles observed for the two ternary complexes relative to the binary complexes (Figs. 1B-D). Moreover, the buried accessible surface at the Sox2/POU_S interface on the *FGF4* enhancer (~240 Å^2) is approximately half that on the *Hoxb1* promoter (~540 Å^2) (7,8) which would predict larger dissociation rate constants for Sox2 and Oct1 in the *FGF4*-DNA ternary complex (see below).

Equilibrium binding of Sox2 and Oct1 to the *FGF4*-DNA was studied by fluorescence anisotropy. The equilibrium dissociation constants for sequence specific binding of Sox2 (K_D^{Sox2}) and Oct1 (K_D^{Oct1}) are 5.3±0.3 and 44±3 nM, respectively, at 150 mM NaCl, 10 mM PIPES, pH 6.5 and 30°C (corresponding exactly to the buffer and temperature conditions used in the NMR experiments) (Fig. 2A). The presence of Sox2 bound to the *FGF4*-DNA duplex increases the sequence specific affinity of Oct1 approximately 15-fold; the equilibrium dissociation constant for sequence specific binding of Oct1 to the Sox2·*FGF4*-DNA binary complex ($K_D^{Oct1\cdot Sox2}$) determined by fluorescence anisotropy is 2.7±0.4 nM (Fig. 2A, *right* panel, filled-in *circles*). Based on the thermodynamic cycle for the binding of Sox2 and Oct1 to DNA (Fig. 2B), the equilibrium dissociation constant for specific DNA binding of Sox2 in the presence of Oct1 ($K_D^{Sox2\cdot Oct1}$) is calculated to be 0.3±0.1 nM (Fig. 2B). By way of comparison, the increase in affinity afforded by protein-

protein interactions within the ternary complex on the *Hoxb1* promoter is approximately 20-fold under slightly different experimental conditions (25°C, 150 mM NaCl, 10 mM phosphate buffer) (18,21).

Global intermolecular translocation of Sox2 and Oct1 on FGF4-DNA — To measure the rate of intermolecular translocation of Sox2 and Oct1 between cognate sites located on different DNA duplexes, we used a similar experimental design to that described in (18,20,21). Single base-pair mutations (Fig. 3A) were introduced in the *FGF4*-DNA duplex (DNA_*a*) adjacent to the recognition sites for Sox2 (DNA_*b*) and the POU_{HD} (DNA_*c*) and POU_S (DNA_*d*) domains of Oct1 (Fig. 3B). These single point mutations have a minimal effect on the binding affinity of Sox2 or Oct1 (18,21), but perturb the $^1H_N/^{15}N$ chemical shifts for a few backbone amide groups of Sox2 or Oct1 in the ternary complexes with DNA_*a* and DNA_*b* or DNA_*c* and DNA_*d* (referred to hereafter as complexes *a* and *b* or *c* and *d*, respectively, Fig. 3C), thereby allowing exchange rates to be measured by monitoring the time-dependence of well-resolved, isolated exchange and auto cross-peaks in a ^{15}N *z*-exchange experiment. In the latter experiment, exchange between $^{15}N_z$ magnetizations from distinct species takes place during the mixing time (that follows the ^{15}N evolution period used to label ^{15}N chemical shifts) giving rise to exchange cross-peaks in a two-dimensional ^1H-^{15}N correlation spectrum (28-30). Exchange rates are obtained by simultaneously fitting the time dependence of the intensities of the exchange and auto cross-peaks as a function of the mixing time using the appropriate McConnell equations (31) for the time development of magnetization in a two-site system (20).

Selective observation of Sox2 or Oct1 was achieved by $[^2H,^{15}N]$-labeling of Sox2 for complexes *a* and *b* or Oct1 for complexes *c* and *d*. Since both Sox2 and Oct1 bind tightly to their cognate DNA sequences with equilibrium dissociation constants in the nanomolar range at 150 mM NaCl (Fig. 2), the ^1H-^{15}N TROSY correlation spectrum for a 1:1 mixture of complexes *a* and *b* (signals from Sox2) or complexes *c* and *d* (signals from Oct1) contains cross-peaks arising from both ternary complexes (18,21). The ratios of the cross-peak intensities for each complex in the ^1H-^{15}N TROSY spectra of the 1:1 mixtures are very close to 1 indicating that the K_D's for both proteins to each DNA duplex are virtually identical.

Examples of *z*-exchange spectra and fits to the time dependence of auto and exchange cross-peak intensities used to determine the exchange rates k_{ab}^{app} and k_{ba}^{app} for Sox2 or k_{cd}^{app} and k_{dc}^{app} for Oct1 are shown in Fig. 3 for His34 of Sox2, and Ser60 and Lys106 of the POU_S and POU_{HD} domains, respectively, of Oct1.

The contributions of jumping and intersegment transfer to intermolecular translocation can be dissected from the dependence of the apparent exchange rate constants on the concentration of free DNA (20,21). The apparent rate constant k_{AB}^{app} for transfer of a protein from site A to site B located on two different DNA molecules is given by the sum of the contributions from jumping and direct intersegment transfer (and similarly for the transfer from sites B to A). With DNA in excess over protein and $k_{off} << k_{on}[DNA_{free}]$, the rate limiting step for jumping is governed by the dissociation rate constant (k_{off}). The jumping rate from A to B is therefore independent of the concentration of free DNA and is given by $k_{off}^{A}/2$, where k_A^{off} is the dissociation rate constant from site A (and the statistical factor of 2 arises from the fact that transfer of the protein between DNA molecules of the same sequence is of equal probability to transfer between DNA molecules of differing sequence). The direct intersegment transfer rate from A to B, on the other hand, is linearly dependent on the free concentration of the DNA containing site B and is given by $k_{AB}^{inter}[DNA_B^{free}]$ where k_{AB}^{inter} is the second order rate constant for direct intersegment transfer from A to B.

k_{ex}^{Sox2} $(= k_{ab}^{app} + k_{ba}^{app})$, and $k_{ex}^{POU_S}$ and $k_{ex}^{POU_{HD}}$ $(= k_{cd}^{app} + k_{dc}^{app}$ for resonances of POU_S and POU_{HD}, respectively) are plotted as a function of free DNA concentration in Fig. 3E. In the context of the *FGF4*-DNA ternary complex, the presence of Oct1 decreases the translocation rates for Sox2 and *vice versa*. Although the exchange rates are slower in the ternary complex than those in the binary complexes, the mechanism of translocation is unaffected by the presence of protein interactions on the DNA. k_{ex}^{Sox2} is independent of the concentration of free DNA indicative of an exclusive jumping mechanism (Fig. 3E, left panel). $k_{ex}^{POU_S}$ (Fig. 3E, middle panel) and $k_{ex}^{POU_{HD}}$ (Fig. 3E, right panel), on the other hand, are linearly dependent on the concentration of free DNA with a measurable

intercept at zero free DNA concentration, indicative of the presence of both direct intersegment transfer and jumping mechanisms.

At 30 ºC, k_{ex}^{Sox2} is reduced from 5.3±0.2 s^{-1} in the binary complex (18) to 1.5±0.3 s^{-1} in the ternary complex (Fig. 3E, left panel). k_{ex}^{Sox2} is equal to the average dissociation rate constant $< k_{off}^{Sox2} >$, since $k_{off}^{a} = 2k_{ab}^{app}$ and $k_{off}^{b} = 2k_{ba}^{app}$ (and note that k_{A}^{off} and k_{B}^{off} in this instance are virtually identical), Given the measured equilibrium dissociation constants for the binding of Sox2 to DNA in the context of binary and ternary complexes (Fig. 2), the average association rate constants $< k_{on}^{Sox2-binary} >$ and $< k_{on}^{Sox2-FGF4\cdot ternary} >$ are calculated to be 1.0(±0.1)x10^9 and 5.1(±1.4)x10^9 $M^{-1}s^{-1}$, respectively (Table 1). Thus, the association rate constant for sequence specific DNA binding of Sox2 is increased approximately 5-fold in the ternary complex relative to the binary one.

The average second-order rate constants for direct intersegment transfer of the POU_S ($< k_{inter}^{POU_S} >$) and POU_{HD} ($< k_{inter}^{POU_{HD}} >$) domains of Oct1 in the ternary complex are 2.2(±0.2)x10^4 and 1.7(±0.1)x10^4 $M^{-1}s^{-1}$, respectively. $< k_{inter}^{POU_S} >$ is reduced by about 50% relative to its value in the binary complex (3.4x10^4 $M^{-1}s^{-1}$ (21)), while $< k_{inter}^{POU_{HD}} >$ remains unaltered (1.8x10^4 $M^{-1}s^{-1}$ in the binary complex (21)). These observations are in complete agreement with the structure of the ternary Sox2·Oct1·*FGF4*-DNA complex (8) since Sox2 interacts only with the POU_S domain of Oct1 (Figs. 1D and 3A). The average dissociation rate constant $< k_{off}^{Oct1} >$ of Oct1 from the ternary complex is 3.5±0.4 s^{-1}, which is comparable to the value of 4.4±0.2 s^{-1} measured for the binary complex (21). The average association rate constant $< k_{on}^{Oct1} >$ for Oct1 binding to the Sox2·*FGF4*-DNA complex is 1.3(±0.2)x10^9 $M^{-1}s^{-1}$, which is approximately 13-fold larger than the value of 1.0(±0.1)x10^8 $M^{-1}s^{-1}$ for the binary complex (21).

The increases in the specific association rate constants for both Sox2 and Oct1 in the context of the ternary complex can probably be attributed to electrostatic enhancement of diffusion controlled association (32,33) afforded by charge-charge interactions between Sox2 and Oct1 when bound specifically to DNA. The approximately three-fold smaller increase in the specific association rate constant for Sox2 compared to Oct1 in the ternary complex can be rationalized as follows. First, the ordering of the C-terminal tail of Sox2 upon interaction with the POU_S domain of Oct1 in the ternary complex (8) (cf. Fig. 3A) entails an entropic penalty since the C-terminal tail is disordered in the binary Sox2·DNA complex (4,18). For the POU_S and POU_{HD} domains of Oct1, ternary complex formation is not accompanied by any significant backbone conformational change, and therefore does not entail any additional entropic cost. (Note that the flexible linker connecting the POU_S and POU_{HD} domains remains largely disordered when bound to DNA (5,7,8)) Second, in the context of the DNA duplexes employed (Fig. 3B) the number of available non-specific sites to which Sox2 can bind and subsequently slide to its specific site is reduced by the presence of Oct1. Although non-specific binding sites available to Oct1 are occluded by the presence of Sox2, there are still a substantial number of non-specific DNA binding sites 3' of the Sox2 binding site (cf. Figs. 3A and B) from which sliding of Oct1 can occur.

Erying plots of the temperature-dependence of the apparent translocation exchange rates (Fig. 3F) provide estimates of the activation enthalpy ($\Delta H^{\ddagger}$), entropy ($T\Delta S^{\ddagger}$ at 30ºC) and by deduction free energy. The activation free energies ($\Delta G^{\ddagger}$) for intermolecular translocation in the ternary and binary complexes on the *FGF4* promoter are comparable (~16-17 kcal·mol^{-1}; Table 2), indicative of similar energy barriers that are unaffected by the presence of a second protein. These data indicate that protein-protein interactions between Sox2 and Oct1 on the DNA modulate translocation rates without perturbing the activation free energies. By inference, this likely holds true for the Sox2·Oct1·*Hoxb1*-DNA ternary complex as well where more extensive protein-protein interactions reduce the translocation rates to levels that are too slow to be measured by *z*-exchange spectroscopy.

Concluding remarks —— The kinetic data on global intermolecular translocation of Oct1 and Sox2 between adjacent specific sites located on different DNA molecules presented here complements our previous work that made use of paramagnetic relaxation enhancement (PRE) measurements (12,15) to examine the interplay between these two transcription factors in translocation events involving sparsely-populated (<1%), highly transient, spectroscopically "invisible" states (18). The latter comprise an ensemble of non-specifically bound species in rapid exchange with the specific complex and participate both in one-

dimensional sliding along the DNA (intramolecular translocation), as well as the formation of bridged intermediates spanning two DNA molecules that precedes intermolecular translocation. The events probed by paramagnetic relaxation enhancement occur on a time scale less than 250-500 μs, although PRE measurements do not afford any further characterization of the kinetics of these processes(12,13,15). Global (or bulk) intermolecular translocation between specific sites on different DNA molecules, on the other hand, occurs on a much slower overall time scale (0.1 to 1 s; cf. Table 1), involves the major spectroscopically visible species (i.e. the specific complexes), and the rate constants from the *z*-exchange experiments pertain directly to the rate-limiting steps in this process (20,21). The interaction of Sox2 and Oct1 on the DNA modulate the translocation mechanisms involving sparsely-populated states (18), as well as the kinetics of global intermolecular translocation between specific sites as shown here. The pathways of global intermolecular translocation, however, are unaffected by the interaction between Sox2 and Oct1.

Based on the kinetic data for global intermolecular translocation presented in this paper, we propose the following model for the sequence of binding, intersegment transfer and dissociation events involved in combinatorial control of gene regulation by Sox2 and Oct1 (Fig. 4). The initial step involves the binding of Sox2 to its specific DNA target site. This is supported by the observation of a ten-fold larger association rate constant for the formation of the binary Sox2·DNA complex versus the Oct1·DNA complex (Table 1), as well as the fact that translocation of Sox2 is a slow process involving only full dissociation (i.e. jumping), whereas Oct1 can undergo rapid global intermolecular translocation at the high DNA concentrations present *in vivo* (~150 mM on a base pair basis). In addition, Sox2 is localized in the cell nucleus (34) while Oct1, which is widely expressed in both adult and embryonic tissues (35,36), is found in both the nucleus and cytoplasm (37). The presence of Sox2 bound to its specific site on the promoter accelerates the binding of Oct1 by ~13-fold to a target site adjacent to the Sox2 site, and the ternary complex is further stabilized by protein-protein interactions, predominantly electrostatic in nature, between Sox2 and Oct1. Once the specific Sox2·Oct1·DNA ternary complex is formed, Oct1 and Sox2 activate transcription synergistically. Oct1 subsequently dissociates from the ternary complex largely via direct intersegment transfer, which, at the high DNA concentrations present *in vivo* will be significantly faster than dissociation into free solution. Intersegment transfer can occur to another specific site on a different promoter or simply to a non-specific site located on either a different DNA molecule or, if on the same DNA, at a widely separated (>150 bp) location through DNA bridging. Finally, Sox2 dissociates from its specific DNA site slowly, and subsequent DNA binding of Sox2 can only occur via a second order reassociation event either to another specific site or to non-specific sites.

Acknowledgments. We thank Drs. Garrett, Baber and Ying for technical support. Y.T. acknowledges a Japan Society for the Promotion of Science Research Fellowship for Japanese Biomedical and Behavioral Researchers at the NIH. This work was supported by the intramural program of NIDDK, NIH and by the AIDS Targeted Antiviral Program of the Office of the Director of the NIH (G.M.C.).

References

1. Wolberger, C. (1999) Multiprotein-DNA complexes in transcriptional regulation. *Ann. Rev. Biophys. Biomol. Struct.* **28**, 29-56
2. Kamachi, Y., Uchikawa, M., and Kondoh, H. (2000) Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet.* **16**, 182-187
3. Dailey, L., and Basilico, C. (2001) Coevolution of HMG domains and homeodomains and the generation of transcriptional regulation by Sox/POU complexes. *J .Cell. Physiol.* **186**, 315-328

4. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G., and Clore, G. M. (2001) Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. *J. Mol. Biol.* **312**, 481-499
5. Klemm, J. D., Rould, M. A., Aurora, R., Herr, W., and Pabo, C. O. (1994) Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**, 21-32
6. Remenyi, A., Tomilin, A., Pohl, E., Lins, K., Philippsen, A., Reinbold, R., Scholer, H. R., and Wilmanns, M. (2001) Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol. Cell. Biol.* **8**, 569-580
7. Williams, D. C., Jr., Cai, M., and Clore, G. M. (2004) Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex. *J. Biol. Chem.* **279**, 1449-1457
8. Remenyi, A., Lins, K., Nissen, L. J., Reinbold, R., Scholer, H. R., and Wilmanns, M. (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* **17**, 2048-2059
9. Berg, O. G., and von Hippel, P. H. (1985) Diffusion-controlled macromolecular interactions. *Annu. Rev. Biophys. Biophys. Chem.* **14**, 131-160
10. von Hippel, P. H., and Berg, O. G. (1989) Facilitated target location in biological systems. *J. Biol. Chem.* **264**, 675-678
11. Halford, S. E., and Marko, J. F. (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* **32**, 3040-3052
12. Clore, G. M., and Iwahara, J. (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* **109**, 4108-4139
13. Clore, G. M. (2011) Exploring sparsely populated states of macromolecules by diamagnetic and paramagnetic NMR relaxation. *Protein Sci.* **20**, 229-246
14. Iwahara, J., Schwieters, C. D., and Clore, G. M. (2004) Characterization of nonspecific protein-DNA interactions by ^{1}H paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **126**, 12800-12808
15. Iwahara, J., and Clore, G. M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227-1230
16. Iwahara, J., Zweckstetter, M., and Clore, G. M. (2006) NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 15062-15067
17. Takayama, Y., and Clore, G. M. (2011) Intra- and intermolecular translocation of the bi-domain transcription factor Oct1 characterized by liquid crystal and paramagnetic NMR. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E169-E176
18. Takayama, Y., and Clore, G. M. (2012) Interplay between the minor and major groove binding transcription factors Sox2 and Oct1 in translocation on DNA studied by paramagnetic and diamagnetic NMR. *J. Biol. Chem.* **287**, 14349-14363.
19. Clore, G. M. (2011) Exploring translocation of proteins on DNA by NMR. *J. Biomol. NMR* **51**, 209-219
20. Iwahara, J., and Clore, G. M. (2006) Direct observation of enhanced translocation of a homeodomain between DNA cognate sites by NMR exchange spectroscopy. *J. Am. Chem. Soc.* **128**, 404-405
21. Doucleff, M., and Clore, G. M. (2008) Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13871-13876
22. Ambrosetti, D. C., Basilico, C., and Dailey, L. (1997) Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions

facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.* **17**, 6321-6329
23. Di Rocco, G., Gavalas, A., Popperl, H., Krumlauf, R., Mavilio, F., and Zappavigna, V. (2001) The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. *J. Biol. Chem.* **276**, 20506-20515
24. Iwahara, J., Anderson, D. E., Murphy, E. C., and Clore, G. M. (2003) EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **125**, 6634-6635
25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293
26. Johnson, B. A., and Blevins, R. A. (1994) NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603-614
27. Sahu, D., Clore, G. M., and Iwahara, J. (2007) TROSY-based *z*-exchange spectroscopy: application to the determination of the activation energy for intermolecular protein translocation between specific sites on different DNA molecules. *J. Am. Chem. Soc.* **129**, 13232-13237
28. Montelione, G. T., and Wagner, G. (1989) 2D chemical exchange NMR spectroscopy by proton-detected heteronuclear correlation. *J. Am. Chem. Soc.* **111**, 3096-3098
29. Perrin, C. L., and Dwyer, T. J. (1990) Application of two-dimensional NMR to kinetics of chemical exchange. *Chem. Rev.* **90**, 935-967
30. Farrow, N. A., Zhang, O. W., Formankay, J. D., and Kay, L. E. (1994) A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical-exchange rates of systems in slow equilibirum. *J. Biomol. NMR* **4**, 727-734
31. McConnell, H. M. (1958) Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430-431
32. Vijayakumar, M., Wong, K. Y., Schreiber, G., Fersht, A. R., Szabo, A., and Zhou, H. X. (1998) Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar. *J. Mol. Biol.* **278**, 1015-1024
33. Zhou, H. X., and Szabo, A. (2004) Enhancement of association rates by nonspecific binding to DNA and cell membranes. *Phys. Rev. Lett* **93**, 178101
34. Zhang, J., Chang, D. Y., Mercado-Uribe, I., and Liu, J. (2012) Sex-determining region Y-box 2 expression predicts poor prognosis in human ovarian carcinoma. *Hum. Pathol. epub ahead of print* doi:10.1016/j.humapath.2011.10.016
35. Kang, J., Shakya, A., and Tantin, D. (2009) Stem cells, stress, metabolism and cancer: a drama in two Octs. *Trends Biochem. Sci.* **34**, 491-499
36. Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D., Sharp, P. A., Ingraham, H. A., Rosenfeld, M. G., Finney, M., Ruvkun, G., and et al. (1988) The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. *Genes Dev.* **2**, 1513-1516
37. Izadpanah, R., Trygg, C., Patel, B., Kriedt, C., Dufour, J., Gimble, J. M., and Bunnell, B. A. (2006) Biologic properties of mesenchymal stem cells derived from bone marrow and adipose tissue. *J. Cell. Biochem.* **99**, 1285-1297

Figure Legends

Fig. 1. Comparison of chemical shift mapping of the protein-protein interaction surfaces at the Sox2 and Oct1 interfaces in the Sox2·Oct1·*FGF4*-DNA and Sox2·Oct1·*Hoxb1*-DNA ternary complexes. *A*, *FGF4* (*left*) and *Hoxb1* (*right*) DNA duplexes. The Sox2 and Oct1 binding sites are delineated by the boxes in *green* and *purple*, respectively. The bases that interact with the POU_S and POU_{HD} domains of Oct1 are indicated by the *red* and *blue* bars, respectively. The *Hoxb1*-DNA duplex represents the actual sequence from the *Hoxb1* promoter (18). The *FGF4*-DNA duplex is not the actual sequence present in the *FGF4* enhancer but simply represents the *Hoxb1* sequence with the three base pair insertion between the Sox2 and Oct1 sites from the *FGF4* enhancer element. As explained in the text, this was done to ensure that differences in equilibrium and kinetic rate constants reflect only the different spacing of the Sox2 and Oct1 sites. *B* and *C*, Profiles of backbone $^1H_N/^{15}N$ chemical shift differences ($\Delta_{H/N}$) between the ternary and binary complexes on the *FGF4*-DNA (*open circles*) and *Hoxb1*-DNA (*light bar*) duplexes for (*B*) Sox2 (*green*) and (*C*) Oct1 (*red*, POU_S domain; *blue*, POU_{HD} domain; *grey*, linker) $\Delta_{H/N}$ is calculated from $[(\Delta\nu_H)^2 + (\Delta\nu_N)^2]^{1/2}$ in Hz at a ^1H frequency of 600 MHz. *D*, Residues showing significant $^1H_N/^{15}N$ chemical shift perturbations mapped onto the structures of the Sox2·Oct1·*FGF4*-DNA (*left*, $\Delta_{H/N} > 20$ Hz, PDB code 1gt0 (8)) and Sox2·Oct1·*Hoxb1*-DNA (*right*, $\Delta_{N/H} > 100$ Hz, PDB code 1o4x (7)) ternary complexes.

Fig. 2. Specific binding of Sox2 to the Oct1·*FGF4*-DNA binary complex monitored by fluorescence anisotropy. *A*, Titration of Sox2 into *FGF4*-DNA (*left panel*) and Oct1 into *FGF4*-DNA (*right panel*, *open circles*) and the Sox2·*FGF4*-DNA complex (*right panel*, filled-in *circles*). The rhodamine fluorescent label was conjugated to the 5' end of the bottom strand of the 32-bp *FGF4*-DNA duplex, and the excitation and emission wavelengths were set to 550 and 580 nm, respectively.. The concentration of the *FGF4*-DNA is 1.5 nM for the Sox2 titration and 10 nM for the Oct1 titration (in the absence and presence of 100 nM Sox2). The temperature is 30°C and the buffer conditions are 150 mM NaCl, 10 mM PIPES, pH 6.5, 94% H2O/6% D_2O, identical to those used in the NMR *z*-exchange experiments. The experimental data points (error bars, 1 s.d.) are displayed as *circles* (open for the binary complexes and filled-in for the ternary complex) and the best-fit curve as *solid (binary complexes) or dashed (ternary complex) lines*. *B*, The equilibrium dissociation constant for Sox2 in the presence of Oct1 was calculated using a thermodynamic cycle and the three experimental K_D values shown.

Fig. 3. Global intermolecular translocation of Sox2 and Oct1 between Sox2·Oct1·*FGF4*-DNA complexes. *A* Ribbon diagram of the crystal structure of the Sox2·Oct1·*FGF4*-DNA complex (PDB code 1gt0 (8)) with Sox2 in *green* and the POU_S and POU_{HD} domains of Oct1 in *red* and *blue*, respectively. The *purple* and *green* meshes indicate the base pair changed in DNA_*a*/DNA_*b* and DNA_*c*/DNA_*d*, respectively. The apparent translocation exchange rates for Sox2, POU_S and POU_{HD} were determined by fitting the time-dependence of the exchange and auto cross-peaks in the *z*-exchange experiment for His34, Ser60, and Lys106, respectively. These three residues exhibit large enough differences in ^1H_N *and* ^{15}N chemical shifts between the DNA_*a* and DNA_*b* complexes (for His34 of Sox2), and between the DNA_*c* and DNA_*d* complexes (for Ser60 and Lys106 of Oct1) to permit accurate quantification of both auto and exchange cross-peak intensities. (*B*) Sequences of *FGF4*-DNA duplexes employed in the *z*-exchange experiments: DNA_*a* and DNA_*b* differ by 1 base pair (*lilac* box) immediately 5' of the Sox2 binding site and were used to measure Sox2 translocation rates; DNA_*c* and DNA_*d*, differ by 2 base pairs (*green* boxes), one immediately 3' of the POU_S site and the other immediately 3' of the POU_{HD} site, and were used to measure Oct1 translocation rates. The Sox2 and Oct1 binding sites are delineated by boxes, and the bases that contact the POU_S and POU_{HD} domains of Oct1 are indicated by the *red* and *blue* bars, respectively. *C*, Examples of *z*-exchange data for His34 of Sox2 (*left*), Ser60 of POU_S (*middle*) and Lys106 of POU_{HD} (*right*) in the Sox2·Oct1·*FGF4*-DNA ternary complex seen in ^1H-^{15}N TROSY-based *z*-exchange spectra at a mixing

time of 220 ms at 30ºC and 150 mM NaCl. For the Sox2 measurements, the concentrations of $^{2}H/^{15}N$-labeled Sox2, Oct1 (at natural isotopic abundance), DNA_*a*, and DNA_*b* are 0.55, 0.7, 0.35, and 0.35 mM, respectively. For the Oct1 measurements, the concentrations of Sox2 (at natural isotopic abundance), $^{2}H/^{15}N$-labeled Oct1, DNA_*c*, and DNA_*d* are 0.7, 0.55, 0.35, and 0.35 mM, respectively. *D*, Time dependence of the intensities of auto (*filled-in circles*) and exchange (*open circles*) cross-peaks at 30ºC and 150 mM NaCl together with the best-fit curves (*solid lines*). The concentrations of proteins and DNA are the same as (*C*). *E* Dependence of the apparent translocation rates for Sox2 (*left*), POU_S (*middle*) and POU_{HD} (*right*) in the ternary complexes (*open circles*) on the concentration of free DNA at 30ºC and 150 mM NaCl. Also shown are the apparent translocation rates for the corresponding binary complexes (solid circles). The protein concentrations are the same as in (*C*) for the ternary complexes; for the binary complexes, the protein at natural isotopic abundance was omitted from the sample. The concentration of free DNA represents the sum of the concentrations of the two DNA duplexes present at a 1:1 ratio in the samples. Association, dissociation and direct intersegment transfer rate constants derived from the data are summarized in Table 1. *F*, Erying plots of the apparent translocation rates at 18, 20, 24, 25, 30, 35 and 40ºC for Sox2 (*left*), POU_S (*middle*) and POU_{HD} (*right*) in ternary (*open circles*) and binary (*filled-in circles*) complexes. The protein and DNA concentrations are the same as (*C*). Linear fits of the apparent translocation rates for the ternary and binary complexes are shown as *solid* and *dashed* lines, respectively. Activation enthalpies, entropies and free energies derived from the data are given in Table 2.

Fig. 4 Model for the sequence of binding, intersegment transfer and dissociation events involved in synergistic transcription regulation by Sox2 and Oct1. The initial event (step *1*) involves the binding of Sox2 to its specific DNA target site, followed by binding of Oct1 to form a specific ternary complex (step 2). The latter binding event is accelerated by the presence of specifically-bound Sox2. Once formation of the ternary Sox2·Oct1-DNA complex has occurred on the promoter or enhancer, transcription of the relevant gene is activated (step *3*). Subsequently, Oct1 dissociates from the DNA largely by direct intersegment transfer to another DNA site (step *4*). Finally, Sox2 dissociates from its specific site into free solution (step *5*). (See text for more details).

Table 1

Intermolecular translocation rates of Sox2 and Oct1 between cognate sites in binary and FGF4-ternary complexes

Intermolecular translocation of Sox2 occurs solely by dissociation followed by reassociation (jumping), while for Oct1 both direct intersegment transfer and jumping occur.

	K_D (nM)	$\langle k_{off} \rangle$ (s^{-1})	$\langle k_{on} \rangle$ ($M^{-1}s^{-1}$)	$\langle k_{inter} \rangle$ ($M^{-1}s^{-1}$)
Sox2				
Binary	5.3±0.3	5.3±0.2	$1.0(\pm0.1)\times10^9$	ND[a]
FGF4-ternary	0.3±0.1	1.5±0.3	$5.1(\pm1.4)\times10^9$	ND[a]
Oct1				
Binary	44±3	4.4±0.2	$1.0(\pm0.1)\times10^8$	$3.4(\pm0.2)\times10^4$ (POU_S) $1.8(\pm0.4)\times10^4$ (POU_{HD})
FGF4-ternary	2.7±0.4	3.5±0.4	$1.3(\pm0.2)\times10^9$	$2.2(\pm0.2)\times10^4$ (POU_S) $1.7(\pm0.1)\times10^4$ (POU_{HD})

[a]Not detectable. From the absence of any concentration dependence in the apparent translocation rates for Sox2, one can conclude that translocation of Sox2 does not involve direct intersegment transfer in both binary and ternary complexes.

Table 2

Apparent activation enthalpies ($\Delta H^{\ddagger}$), entropies ($\Delta S^{\ddagger}$), and free energies ($\Delta G^{\ddagger}$) for global intermolecular translocation of Sox2 and Oct1 between cognate sites in the binary and FGF4-ternary complexes

	$\Delta H^{\ddagger}$ (kcal·mol^{-1})[a]	$\Delta S^{\ddagger}$ (cal·mol^{-1}·K^{-1})[a]	$\Delta G^{\ddagger}$ (kcal·mol^{-1})[b]
Sox2			
Binary / *FGF4*-ternary	21.2±1.9 / 21.3±1.8	15.0±6.5 / 13.1±5.8	16.6±2.8 / 17.3±2.5
Oct1-POU_S domain			
Binary / *FGF4*-ternary	23.1±1.8 / 20.4±0.8	22.6±4.0 / 12.7±2.4	16.3±2.2 / 16.5±1.1
Oct1-POU_{HD} domain			
Binary / *FGF4*-ternary	20.8±1.2 / 20.3±0.5	14.7±4.1 / 12.7±1.6	16.4±1.7 / 16.5±0.7

[a]Values were derived by least squares fitting of the Erying plots shown in Fig. 3F.

[b]$\Delta G^{\ddagger}$ is calculated from $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$ at 303 K.

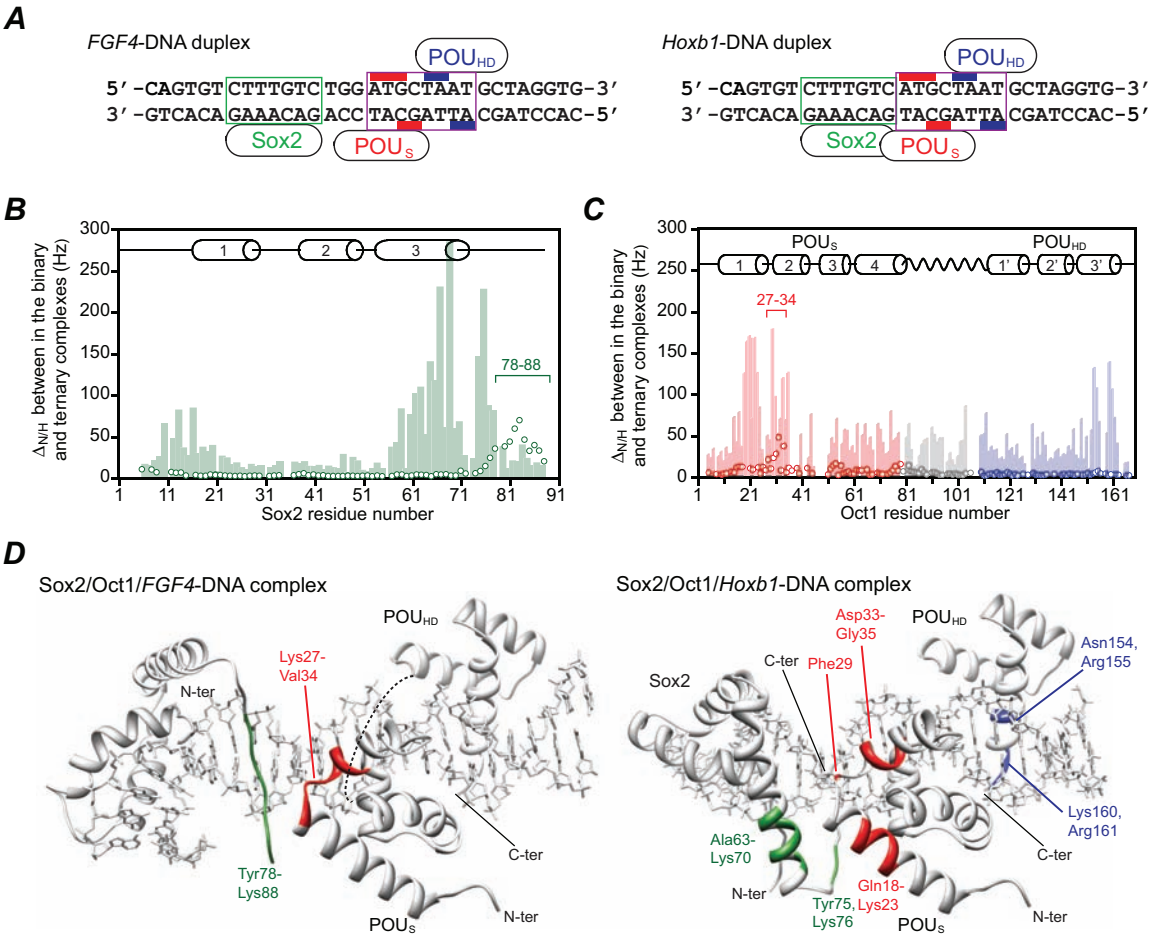

Fig. 1

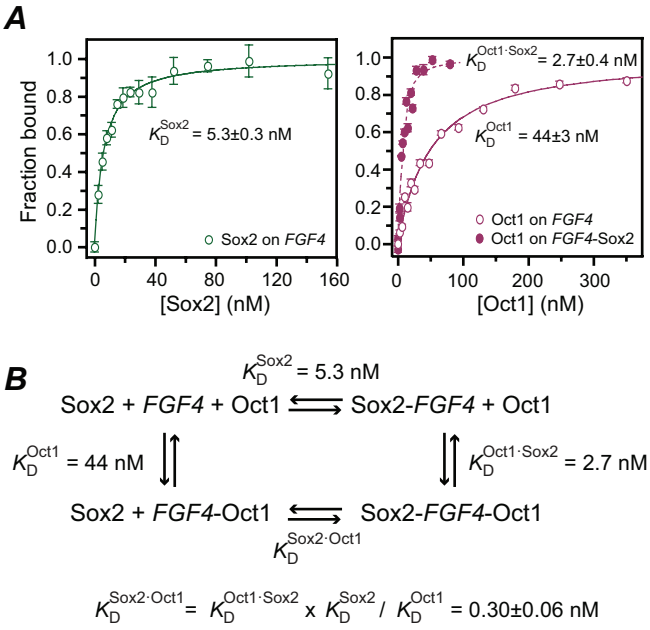

Fig. 2

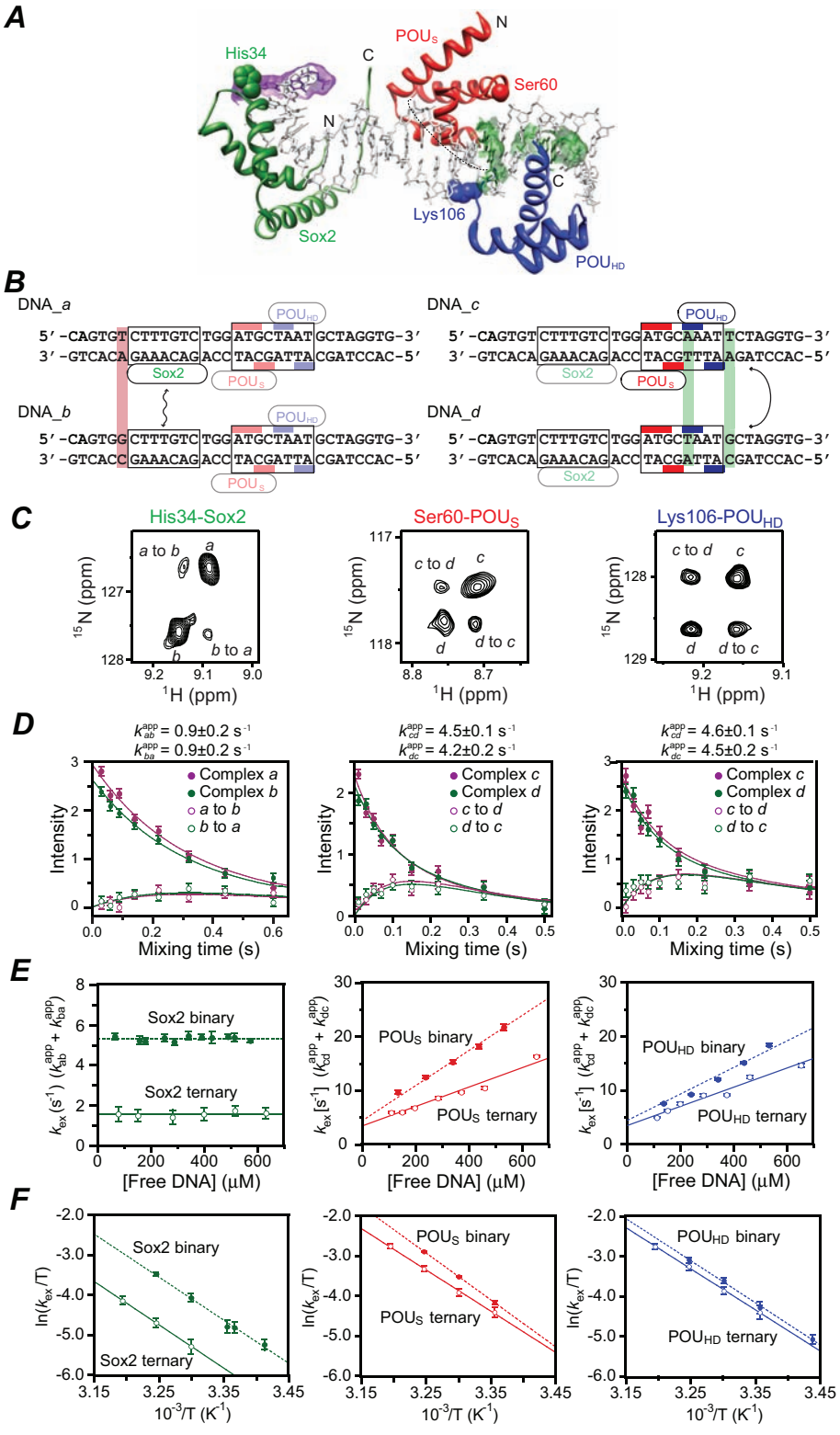

Fig. 3

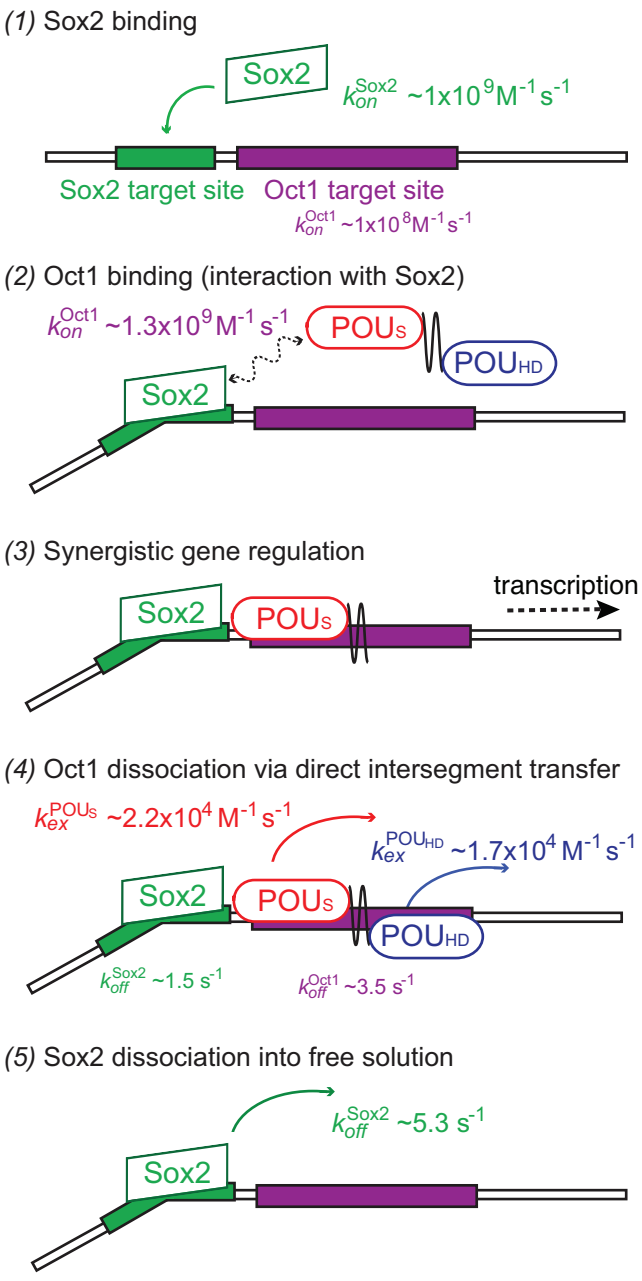

Fig. 4